

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2003, 21:14:21 ; Search time 102 Seconds  
(without alignments)  
1214.365 Million cell updates/sec

Title: US-09-830-980-1  
Perfect score: 2574  
Sequence: 1 MQETDTEQEATPHTIQARLV.....PDGSRVASGGKDKVIKLWAY 480

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2566	99.7	480	5	O96995
2	2554.5	99.2	481	5	Q9VPR4
3	2554.5	99.2	488	5	Q8T4A2
4	1520	59.1	485	4	Q9NVX2
5	1520	59.1	487	4	Q9BU54
6	1513	58.8	476	13	O93531
7	1512	58.7	485	11	Q8VEJ4
8	1335.5	51.9	473	10	Q9FLX9
9	1331	51.7	447	10	Q9AYI7
10	1318.5	51.2	473	5	Q9TYV3
11	1093	42.5	515	3	Q8NKJ4
12	1047.5	40.7	502	3	O74855
13	834.5	32.4	645	5	Q8IHQ8
14	616	23.9	1356	3	Q8X1P4
15	603	23.4	1356	3	Q8X1P3
16	603	23.4	1356	3	Q8X1P5

17	595	23.1	1376	3	Q8X1P2	Q8xlp2 podospora a
18	538.5	20.9	1708	16	Q8YZI2	Q8yzi2 anabaena sp
19	533	20.7	1227	16	Q8Z0R1	Q8z0r1 anabaena sp
20	532.5	20.7	1711	16	Q8Z0I9	Q8z0i9 anabaena sp
21	532	20.7	1747	16	Q8Z0Z0	Q8z0z0 anabaena sp
22	497	19.3	934	16	Q8YZZ3	Q8yzz3 anabaena sp
23	496.5	19.3	1189	16	Q8YL09	Q8yl09 anabaena sp
24	486	18.9	1551	16	Q8YMU3	Q8ymu3 anabaena sp
25	476.5	18.5	676	16	Q8YSG6	Q8ysg6 anabaena sp
26	469	18.2	1189	16	Q8YTD1	Q8ytd1 anabaena sp
27	462.5	18.0	1241	2	Q9XBD8	Q9xbd8 amycolatops
28	451	17.5	677	16	Q8YZL9	Q8yzl9 anabaena sp
29	436	16.9	304	16	Q8Z0S4	Q8z0s4 anabaena sp
30	425.5	16.5	1049	2	Q9ZEM4	Q9zem4 streptomyce
31	425.5	16.5	1676	16	Q9KXX9	Q9kxx9 streptomyce
32	425	16.5	329	16	Q8XB12	Q8xb12 chlorobium
33	424.5	16.5	559	16	Q8YSC0	Q8ysc0 anabaena sp
34	412.5	16.0	700	5	Q9XZK1	Q9xzk1 drosophila
35	411	16.0	353	5	Q8T776	Q8t776 brachiosto
36	410.5	15.9	317	10	Q9M2Z2	Q9m2z2 arabidopsis
37	400	15.5	589	16	Q8YN14	Q8ynl4 anabaena sp
38	399.5	15.5	414	5	Q8SW59	Q8sw59 encephalito
39	399	15.5	330	4	Q9NUL4	Q9nul4 homo sapien
40	396	15.4	613	10	Q9FN19	Q9fnl9 arabidopsis
41	389	15.1	411	5	O96698	O96698 drosophila
42	386.5	15.0	1051	17	Q8TMS3	Q8tm83 methanosarc
43	385.5	15.0	328	11	Q9D7H2	Q9d7h2 mus musculu
44	385.5	15.0	415	4	Q8N136	Q8nl36 homo sapien
45	377.5	14.7	415	4	Q8N776	Q8n776 homo sapien

ALIGNMENTS

RESULT 1

O96995	ID	O96995	PRELIMINARY;	PRT;	480 AA.
AC	O96995;				
DT	01-MAY-1999	(TREMBlrel. 10, Created)			
DT	01-MAY-1999	(TREMBlrel. 10, Last sequence update)			
DT	01-OCT-2002	(TREMBlrel. 22, Last annotation update)			
DE	NOTCHLESS	protein.			
GN	NLE OR NOTCHLESS	OR CG2863.			
OS	Drosophila melanogaster	(Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99077802; PubMed=9857191;				
RA	Royet J., Bouwmeester T., Cohen S.M.;				
RT	*Notchless encodes a novel WD40-repeat-containing protein that				
RT	modulates Notch signaling activity."				
RL	EMBO J. 17:7351-7360(1998).				
DR	EMBL; AJ012588; CAA10070.1; --				
DR	FlyBase; FBgn0021874; Nle.				
DR	InterPro; IPR001632; Gprotein_B.				
DR	InterPro; IPR001680; WD40.				
DR	Pfam; PF00400; WD40; 8.				
DR	PRINTS; PR00319; GPROTEINB.				
DR	PRINTS; PR00320; GPROTEINRPT.				
DR	ProDom; PD000018; WD40; 6.				
DR	SMART; SM00320; WD40; 8.				
DR	PROSITE; PS00678; WD_REPEATS_1; 2.				
DR	PROSITE; PS50082; WD_REPEATS_2; 7.				
DR	PROSITE; PS50294; WD_REPEATS_REGION; 1.				
KW	Repeat; WD repeat.				
SQ	SEQUENCE 480 AA; 52933 MW; 40A5D696D33956C4 CRC64;				

Query Match 99.7%; Score 2566; DB 5; Length 480;  
Best Local Similarity 99.8%; Pred. No. 1.2e-213;  
Matches 479; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MQETDTEQATPHTIOARLVYTGEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLFF 60  
Db 1 MQETDTEQATPHTIOARLVYTGEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLFF 60  
QY 61 VGEDEIKKSLEDTLDLASVDNTENVIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNFS 120  
Db 61 VGEDEIKKSLEDTLDLASVDNTENVIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNFS 120  
QY 121 PDGAHLASGSGDITVRLWDLNTETPHFTCTGCHKQWVLCVSWAPDGKRLASGCKAGSIIV 180  
Db 121 PDGAHLASGSGDITVRLWDLNTETPHFTCTGCHKQWVLCVSWAPDGKRLASGCKAGSIIV 180  
QY 181 DPETGQOKGRPLSGHKHCHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 240  
Db 181 DPETGQOKGRPLSGHKHCHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 240  
QY 241 GHTNAVAVRWGGAGLIYTSSKDRITVCMWRAADGILCRTHSGHAHVVNNIALSTDYVLR 300  
Db 241 GHTNAVAVRWGGAGLIYTSSKDRITVCMWRAADGILCRTHSGHAHVVNNIALSTDYVLR 300  
QY 301 GPFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCDDNTLYLWRNNQKCV 360  
Db 301 GPFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCDDNTLYLWRNNQKCV 360  
QY 361 ERMTHGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSAD 420  
Db 361 ERMTHGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSAD 420  
QY 421 SRLIVSGSKDSTLKVMSVQTKKLAQELPGHAEVFGVDWAPDGSRVASGGKDKVIKLWAY 480  
Db 421 SRLIVSGSKDSTLKVMSVQTKKLAQELPGHAEVFGVDWAPDGSRVASGGKDKVIKLWAY 480

RESULT 2

Q9VPR4 PRELIMINARY; PRT; 481 AA.  
AC Q9VPR4;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE NLE protein.  
GN NLE OR CG2863.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulo G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).  
DR EMBL; AB003589; AAF51479.1; -.  
DR FlyBase; FBgn0021874; Nle.  
DR InterPro; IPR001632; GproteineB.  
DR InterPro; IPR001680; WD40.  
DR Pfam; PF00400; WD40; 9.  
DR PRINTS; PR00319; GPROTEINB.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PRODom; PD000018; WD40; 6.  
DR SMART; SM00320; WD40; 8.  
DR PROSITE; PS00678; WD\_REPEATS\_1; 2.  
DR PROSITE; PS0082; WD\_REPEATS\_2; 7.  
DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 481 AA; 52981 MW; FOEDD0815E691914 CRC64;

Query Match 99.2%; Score 2554.5; DB 5; Length 481;  
Best Local Similarity 99.6%; Pred. No. 1.2e-212;  
Matches 479; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MQETDTEQATPHTIOARLVY-TGEEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLF 59  
Db 1 MQETDTEQATPHTIOARLVSDTGEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLF 60  
QY 60 FVGEDEIKKSLEDTLDLASVDNTENVIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNF 119  
Db 61 FVGEDEIKKSLEDTLDLASVDNTENVIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNF 120  
QY 120 SPDGAHLASGSGDITVRLWDLNTETPHFTCTGCHKQWVLCVSWAPDGKRLASGCKAGSI 179  
Db 121 SPDGAHLASGSGDITVRLWDLNTETPHFTCTGCHKQWVLCVSWAPDGKRLASGCKAGSI 180  
QY 180 WDPETGQOKGRPLSGHKHCHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 239  
Db 181 WDPETGQOKGRPLSGHKHCHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 240  
QY 240 AGHTNAVAVRWGGAGLIYTSSKDRITVCMWRAADGILCRTHSGHAHVVNNIALSTDYVLR 299  
Db 241 AGHTNAVAVRWGGAGLIYTSSKDRITVCMWRAADGILCRTHSGHAHVVNNIALSTDYVLR 300  
QY 300 TGPFPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCDDNTLYLWRNNQK 359  
Db 301 TGPFPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCDDNTLYLWRNNQK 360  
QY 360 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSA 419  
Db 361 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSA 420  
QY 420 DSRLIVSGSKDSTLKVMSVQTKKLAQELPGHAEVFGVDWAPDGSRVASGGKDKVIKLWA 479  
Db 421 DSRLIVSGSKDSTLKVMSVQTKKLAQELPGHAEVFGVDWAPDGSRVASGGKDKVIKLWA 480  
QY 480 Y 480  
Db 481 Y 481

RESULT 3  
Q8T4A2 PRELIMINARY; PRT; 488 AA.  
ID Q8T4A2;  
AC Q8T4A2;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE AT08344p.  
GN NLE OR CG2863.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY089286; AAL90024.1; -.  
DR FlyBase; FBgn0021874; Nle.  
DR InterPro; IPR001680; WD40.  
DR Pfam; PF00400; WD40; 6.  
DR ProDom; PD000018; WD40; 8.  
DR SMART; SM00320; WD40; 8.  
DR PROSITE; PS00678; WD\_REPEATS\_1; 2.  
DR PROSITE; PS0082; WD\_REPEATS\_2; 7.  
DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 488 AA; 53809 MW; D4883202954A7785 CRC64;  
  
Query Match 99.2%; Score 2554.5; DB 5; Length 488;  
Best Local Similarity 99.6%; Pred. No. 1.2e-212;  
Matches 479; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
  
QY 1 MQETDTEQATPHTIQARLVY-TGEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLF 59  
DB 8 MQETDTEQATPHTIQARLVSDTGEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLF 67  
  
QY 60 FVGEDEIKKSLEDTLDLASVDTENVIDIVYQQAQVFKVRPVTCTSSMPGHAEAVVSLNF 119  
DB 68 FVGEDEIKKSLEDTLDLASVDTENVIDIVYQQAQVFKVRPVTCTSSMPGHAEAVVSLNF 127  
  
QY 120 SPDGAHLASGSGDTTVRLWDLNLTETPHFTCTGKHQWVLCVSWAPDGRKRLASGCKAGSI 179  
DB 128 SPDGAHLASGSGDTTVRLWDLNLTETPHFTCTGKHQWVLCVSWAPDGRKRLASGCKAGSI 187  
  
QY 180 WDPETGQKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 239  
DB 188 WDPETGQKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 247  
  
QY 240 AGHTNAVTAVRWGGAGLIYTSSKDRITVKMWRAADGILCRFTSGHAHWNNIALSTDYVLR 299  
DB 248 AGHTNAVTAVRWGGAGLIYTSSKDRITVKMWRAADGILCRFTSGHAHWNNIALSTDYVLR 307  
  
QY 300 TGFPHVPKDRSKSHLSLSTEELQESALKRYQAVCPDEVESSLVSCSDDNTLYLWRNNQNC 359  
DB 308 TGFPHVPKDRSKSHLSLSTEELQESALKRYQAVCPDEVESSLVSCSDDNTLYLWRNNQNC 367  
  
QY 360 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSA 419  
DB 368 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSA 427  
  
QY 420 DSRLLVSGSKDSTLKVMSVQTKLAQELPGHADEVFQVWAPDGRSVASGGKDKVILKWA 479  
DB 428 DSRLLVSGSKDSTLKVMSVQTKLAQELPGHADEVFQVWAPDGRSVASGGKDKVILKWA 487  
  
QY 480 Y 480

Db 488 Y 488  
  
RESULT 4  
Q9NVX2 PRELIMINARY; PRT; 485 AA.  
ID Q9NVX2;  
AC Q9NVX2;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE Hypothetical protein FLJ10458.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.;  
RT "NEDO human cDNA sequencing project.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Strausberg R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK001320; BAA91621.1; -.  
DR EMBL; BC012075; BAA12075.1; -.  
DR InterPro; IPR001632; Gproteins\_B.  
DR InterPro; IPR001680; WD40.  
DR Pfam; PF00400; WD40; 8.  
DR PRINTS; PR00319; GPROTEINB.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR ProDom; PD000018; WD40; 7.  
DR SMART; SM00320; WD40; 8.  
DR PROSITE; PS00678; WD\_REPEATS\_1; 4.  
DR PROSITE; PS0082; WD\_REPEATS\_2; 7.  
DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
KW Hypothetical protein; Repeat; WD repeat.  
SQ SEQUENCE 485 AA; 53266 MW; 0AEBE24B44957379 CRC64;  
  
Query Match 59.1%; Score 1520; DB 4; Length 485;  
Best Local Similarity 58.6%; Pred. No. 5.4e-123;  
Matches 279; Conservative 73; Mismatches 118; Indels 6; Gaps 3;  
  
QY 9 EATPHTIQARLVYTGEEA---GPPIDLPAGITTOQLGLICNALLKNEEATPYLFFVGED 64  
DB 8 EAVARDVQRLVQFQDEGGQLLGSFPDVPDITPDRQLQVNCNALLAQEDPLPLAFFVHDA 67  
  
QY 65 EIKKSLEDTLDLASVDTENVIDIVYQQAQVFKVRPVTCTSSMPGHAEAVVSLNFSPDGA 124  
DB 68 EIVSSLGKTLESQAQVETEKVLDIIYQQAIFRVRVAVTRCTSSLEGHSEAVISVAFSPTGK 127  
  
QY 125 HLASGSGDTTVRLWDLNLTETPHFTCTGKHQWVLCVSWAPDGRKRLASGCKAGSI 184  
DB 128 YLASGSGDTTVRFWDLSTETPHFTCTGKHHRHVLSISWSPDGKRLASGCKNGQILLWDPST 187  
  
QY 185 GQKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNIAGHTN 244  
DB 188 GKQVGRTLAGHSKWITGLSWEPLHANPECRYVASSSKOGSVRIWDTTAGRCERILTGHTQ 247  
  
QY 245 AVTAVRWGGAGLIYTSSKDRITVKMWRAADGILCRFTSGHAHWNNIALSTDYVLRTPFH 304  
DB 248 SVTCLRWGGDGLYSASQDRTIKVWRAHDGVLCTLQGHGHWNTMALSTDYALRTGAFE 307  
  
QY 305 PVK-DRSKSHLSLSTEELQESALKRYQAVCPDEVESSLVSCSDDNTLYLWRNNQN-KCVER 362  
DB 308 PAEASVNPQDLQSLQELKERALSRYNLVRGQGPRLVSGSDDFTLFLWSPAEDKKPLTR 367





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2003, 21:07:26 ; Search time 23 Seconds  
(without alignments)  
981.427 Million cell updates/sec

Title: US-09-830-980-1  
Perfect score: 2574  
Sequence: 1 MQETDTEQATPHTIQARLV.....PDGSRVASGGKDKVIKLWAY 480

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1092	42.4	515	1 YCW2 YEAST	P25382 saccharomyc
2	599	23.3	1258	1 YS00 ANASP	Q8ytc2 anabaena sp
3	597	23.2	1683	1 YL24 ANASP	Q8yv57 anabaena sp
4	590	22.9	1526	1 YV46 ANASP	Q8yr11 anabaena sp
5	581	22.6	1356	1 HET1 PODAN	Q00808 podospora a
6	433	16.8	1693	1 Y163 SYNY3	Q55563 synecocyst
7	419	16.3	334	1 WDR5 HUMAN	Q9ugp9 homo sapien
8	415	16.1	361	1 WDS DROME	Q9v3j8 drosophila
9	409.5	15.9	742	1 PKWA THECU	P49695 thermomonos
10	377	14.6	522	1 TBLR HUMAN	Q9bq87 homo sapien
11	372.5	14.5	514	1 TBLR HUMAN	Q9bzk7 homo sapien
12	372	14.5	376	1 YKY4 CAEEL	Q17963 caenorhabdi
13	371	14.4	526	1 TBLX HUMAN	Q60907 homo sapien
14	361.5	14.0	501	1 YH92 CAEEL	Q23256 caenorhabdi
15	359	13.9	614	1 TUI1 SCHPO	Q09715 schizosacch
16	354	13.8	409	1 LIS1 BOVIN	P43033 bos taurus
17	354	13.8	409	1 LIS1 HUMAN	P43034 homo sapien
18	354	13.8	409	1 LIS1 MOUSE	P43035 mus musculu
19	347.5	13.5	395	1 YZLL CAEEL	Q93847 caenorhabdi
20	346	13.4	678	1 SCOB EMENI	Q00659 emericeella
21	320	12.4	423	1 WD12 HUMAN	Q9gzl7 homo sapien
22	319.5	12.4	542	1 FW1B HUMAN	Q9ukb1 homo sapien
23	319	12.4	312	1 GBLP LEIMA	Q25306 leishmania
24	318.5	12.4	586	1 TUI2 SCHPO	Q9uug8 schizosacch
25	316	12.3	423	1 WD12 MOUSE	Q9jjja4 mus musculu
26	315	12.2	312	1 GBLP LEICH	Q27434 leishmania
27	314.5	12.2	518	1 TRCB XENLA	Q91854 xenopus lae
28	313.5	12.2	605	1 FW1A HUMAN	Q9y297 homo sapien
29	312.5	12.1	605	1 POF1 SCHPO	P87053 schizosacch
30	312.5	12.1	665	1 LI23 CAEEL	Q09990 caenorhabdi
31	310.5	12.1	640	1 MT30 YEAST	P39014 saccharomyc
32	309	12.0	608	1 WDR1 DROME	Q9vu68 drosophila
33	305	11.8	514	1 TUP1 CANAL	P56093 candida alb

34	303.5	11.8	682	1 TUP1 KLULA	P56094 kluyveromyc
35	303	11.8	1248	1 APAF HUMAN	O14727 homo sapien
36	303	11.8	1249	1 APAF RAT	Q9epv5 rattus norv
37	301	11.7	606	1 PF20 CHURE	P93107 chlamydomon
38	300	11.7	604	1 RCO1 NEUCR	P78706 neurospora
39	300	11.7	1249	1 APAF MOUSE	O88879 mus musculu
40	297	11.5	579	1 SE10 CAEEL	Q93794 caenorhabdi
41	296	11.5	340	1 GBB1 CAEEL	P17343 caenorhabdi
42	296	11.5	609	1 WDR1 CHICK	O93277 gallus galli
43	294.5	11.4	713	1 TUP1 YEAST	P16649 saccharomyc
44	292.5	11.4	404	1 WSB2 HUMAN	Q9ny87 homo sapien
45	291	11.3	704	1 T2D4 DROME	P49846 drosophila

ALIGNMENTS

RESULT 1  
YCW2 YEAST  
ID YCW2 YEAST STANDARD; PRT; 515 AA.  
AC P25382;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical 57.0 kDa Trp-Asp repeats containing protein in CPR4-SSK22  
DE intergenic region.  
GN YCR072C OR YCR72C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ballesta J.P.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,  
RA Sanz E.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [2]

REVISIONS.  
RA Jimenez A.;  
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 8 WD repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; X59720; CAA42270.1; -.  
CC PIR; S19487; S19487.  
CC SGD; S0000668; YCR072C.  
CC InterPro; IPR001632; Gpotein\_B.  
CC InterPro; IPR001680; WD40.  
CC Pfam; PF00400; WD40; 8.  
CC PRINTS; PR00319; GPROTEINB.  
CC PRINTS; PR00320; GPROTEINBRPT.  
CC ProDom; PD000018; WD40; 6.  
CC SMART; SM00320; WD40; 8.  
CC PROSITE; PS00678; WD\_REPEATS\_1; 5.  
CC PROSITE; PS50082; WD\_REPEATS\_2; 7.  
CC PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
KW Hypothetical protein; Repeat; WD repeat.  
FT REPEAT 141 172  
FT REPEAT 184 214  
FT REPEAT 227 264  
FT REPEAT 276 305  
FT REPEAT 318 387  
FT REPEAT 400 430  
FT REPEAT 442 472  
FT REPEAT 484 514  
SQ SEQUENCE 515 AA; 57040 MW; DFE603B79BFB530A CRC64;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2003, 21:14:46 ; Search time 43 Seconds  
(without alignments)  
1073.511 Million cell updates/sec

Title: US-09-830-980-1  
Perfect score: 2574  
Sequence: 1 MQETDTEQATPHTIQARLV.....PDGSRVASGGKDKVIXLMAY 480

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1318.5	51.2	473	2 T33805	hypothetical prote
2	1092	42.4	515	2 S19487	hypothetical prote
3	1047.5	40.7	502	2 T41148	trp-aap repeat con
4	599	23.3	1258	2 AI2155	WD-repeat protein
5	597	23.2	1683	2 AF2071	WD-40 repeat prote
6	590	22.9	1526	2 AC2239	WD-40 repeat prote
7	581	22.6	1356	2 T18521	beta transducin-li
8	538.5	20.9	1708	2 AE1866	WD-40 repeat prote
9	533	20.7	1227	2 AE1810	WD-40 repeat prote
10	532.5	20.7	1711	2 AD1842	WD-40 repeat prote
11	532	20.7	1747	2 AC1842	WD-40 repeat prote
12	497	19.3	934	2 AG1889	WD-40 repeat prote
13	496.5	19.3	1189	2 AI2493	WD-repeat protein
14	486	18.9	1551	2 AB2410	WD-repeat protein
15	476.5	18.5	676	2 AH2195	hypothetical prote
16	469	18.2	1189	2 AH2154	WD-repeat protein
17	451	17.5	677	2 AE1861	serine/threonine k
18	436	16.9	304	2 AG1837	WD-40 repeat prote
19	433	16.8	1693	2 S76086	beta transducin-li
20	425.5	16.5	1049	2 T42045	beta transducin-li
21	424.5	16.5	559	2 AB2202	hypothetical prote
22	410.5	15.9	317	2 T46032	WD-40 repeat regul
23	400	15.5	589	2 AG2400	WD-repeat protein
24	372	14.5	376	2 T19266	hypothetical prote
25	371	14.4	786	2 AG2375	WD-40 repeat-prote
26	368.5	14.3	777	2 T41075	hypothetical WD-re
27	365.5	14.2	876	2 T51507	WD40-repeat protei
28	361.5	14.0	501	2 T27513	hypothetical prote
29	359	13.9	614	2 S58306	WD-40 repeat regul

30	355	13.8	333	2 G85034	probable WD-repeat
31	354	13.8	409	2 S36113	LIS-1 protein - hu
32	354	13.8	410	2 S48052	platelet-activatin
33	347.5	13.5	395	2 T23317	hypothetical prote
34	344.5	13.4	265	2 AF1890	WD-repeat protein
35	327.5	12.7	357	2 AI2099	WD-40 repeat prote
36	323.5	12.6	343	2 C84870	probable splicing
37	322	12.5	342	2 AE2490	WD-repeat protein
38	318.5	12.4	586	2 T38992	WD-40 repeat regul
39	314.5	12.2	518	2 B48088	beta-transducin re
40	314.5	12.2	598	2 AE2415	WD-repeat protein
41	312.5	12.1	323	2 T02617	hypothetical prote
42	312.5	12.1	605	2 T38932	probable sulfur me
43	312.5	12.1	701	2 T16607	hypothetical prote
44	312	12.1	1194	2 T03818	apoptotic proteina
45	310.5	12.1	640	2 S49932	MEF30 protein - ye

ALIGNMENTS

RESULT 1

T33805  
hypothetical protein W07E6.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 26-May-2000  
C:Accession: T33805  
R:Latreille, P.; Wamsley, P.  
submitted to the EMBL Data Library, November 1998  
A:Description: The sequence of C. elegans cosmid W07E6.  
A:Reference number: Z21414  
A:Accession: T33805  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-473 <LAT>  
A:Cross-references: EMBL:AF106576; PIDN:AAC78176.1; GSPDB:GN00020; CESP:W07E6.2  
A:Experimental source: strain Bristol N2; clone W07E6  
C:Genetics:  
A:Gene: CESP:W07E6.2  
A:Map position: 2  
A:Introns: 46/2; 77/3; 103/1; 195/2; 256/3; 311/3; 399/2  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match	51.2%	Score	1318.5;	DB	2;	Length	473;
Best Local Similarity	53.8%	Pred. No.	9e-95;				
Matches	250;	Conservative	72;	Mismatches	122;	Indels	21;
Gaps	7;						
Qy	24	EEAGPPIDLPA	GTTQQLGLICNALL-KNEBATPYLFFVGE-----DEIKKSLEDTDLA	77			
Db	18	ELGGGILVPVD	ISTNELQILCNQLGSSDDPVPISFFTEGAEIVDSIRKSLEE-----	72			
Qy	78	SVDTENVIDIV	YQQAQVFKVRPVTRCTSSMPGHABAVVSLNFSPDGAHLASGSGDITVRL	137			
Db	73	-IDFETTLKLV	YQQAQVFRVRPVTRCSASIPGHGEPVISAQFSPDGRGLASGSGDQTMRI	131			
Qy	138	WDLNLTETPFT	CTGHKQWVLCVSWAPDGRKRLASGCKAGSIIIWDPETGQQKGRPLSGHKK	197			
Db	132	WDIELEPLHTC	CKSHKSWVLCIAWSPDATKIASACNGEICIWNNAKTGEQIGKTLKRHKQ	191			
Qy	198	HINCLAWEPYH	RDECRKLASASGDGDCRIWDVKLGQCLMNIAGHTNAVAVRWGGAGLI	257			
Db	192	WITSLAWQPMH	KDPTCRLLASCGKGNIFWDTVQGTVVRCLSGHTASVTCLRWGGEGLI	251			
Qy	258	YTSSKDRTVKM	WRAADGILCRFTSGHAHVNNIALSTDYVLRTPGPHPVKDRSKSHLSLS	317			
Db	252	YSGQDRTVKM	WRADGVMCRNMTGHAHWINTLALNTDYALRTSCFEPSPKRCIKPD---T	308			
Qy	318	TEELQESALKR	YQAVCPDEV--ESLVSCSDNTLYLWRNNQNK-CVERMTGHQNVNDV	373			
Db	309	VERCQKVAQTR	YEAL--EIAGGERLVSGSDFTLFMWNPKETKQSNRMVTHMQLVNVQ	366			
Qy	374	KYSPDVKLIAS	ASFDKSVRLWRASDGQYMATFRGHVQAVYTVVWASDSRSLIVSGSKDSTL	433			









[illegible]

```

RESULT 2
US-10-132-744A-2
; Sequence 2, Application US/10132744A
; Publication No. US20030027261A1
; GENERAL INFORMATION:
; APPLICANT: Utku, Nalan
; TITLE OF INVENTION: No. US20030027261A1el genes Tzap7/A, Tzap7/B and Tzap7 involved in
; TITLE OF INVENTION: activation and uses thereof
; FILE REFERENCE: Utku-4 CON
; CURRENT APPLICATION NUMBER: US/10/132,744A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: PCT/EP00/10670
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 60/185,016
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/162,675
; PRIOR FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 484
; TYPE: prt
; ORGANISM: Homo sapiens
; US-10-132-744A-2

```

QY	245	AVTAVRWGGAGLIYTSSKORTVTKMWRADGILCRFTFSGHAHWNNIALSTDYVLRGPFH	304
Db	247	SVTCLRWGGDGLLYSASQDRTIKWRAHDGVLCTLQGHGHVWNTMALSTDYALRTGAFE	306
QY	305	PVK-DRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDDNTLYLWRNNQN-KCVER	362
Db	307	PAEASVNPQDLQGSLSQELKERALSRYNLVRGQGPRLVSGSDDFTLFLWSPAEDKKPLTR	366
QY	363	MTGHQNVNDVKYSPDKLIASAFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSADSR	422
Db	367	MTGHQALINQVLFSPDSRIVASAFDKSIKLWDGRTGKYLASLRGHVAAVYQIAWSADSR	426
QY	423	LIVSGSKDSTLKVMSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVTKLW	478
Db	427	LLVSGSSDSTLKVMDVKAQKLA MDLPGHADEVYAVDWSPDGQORVASGGKDKCLRIW	482

RESULT 3  
 US-10-128-714-8213  
 ; Sequence 8213, Application US/10128714  
 ; Publication No. US20030119013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Bo  
 ; APPLICANT: Hu, Wenqi  
 ; APPLICANT: Tishkoff, Daniel  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Eroshkin, Alexey M  
 ; APPLICANT: Lemieux, Sebastien M  
 ; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus*  
 ; TITLE OF INVENTION: Methods of Use  
 ; FILE REFERENCE: 10182-018-999  
 ; CURRENT APPLICATION NUMBER: US/10/128,714  
 ; CURRENT FILING DATE: 2002-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/285,697  
 ; PRIOR FILING DATE: 2001-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/287,066  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/295,890  
 ; PRIOR FILING DATE: 2001-06-05  
 ; PRIOR APPLICATION NUMBER: US 60/303,899  
 ; PRIOR FILING DATE: 2001-07-09  
 ; PRIOR APPLICATION NUMBER: US 60/316,362  
 ; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 8213  
 ; LENGTH: 515  
 ; TYPE: PRT  
 ; ORGANISM: *Aspergillus fumigatus*  
 US-10-128-714-8213

[illegible]

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2003, 21:16:01 ; Search time 29 Seconds  
(without alignments)  
700.318 Million cell updates/sec

Title: US-09-830-980-1  
Perfect score: 2574  
Sequence: 1 MQETDTEQATPHTIQARLV.....PDGSRVASGGKDKVIKLWAY 480

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1058.5	41.1	514	1 US-08-190-802A-66	Sequence 66, Appl
2	1058.5	41.1	514	3 US-08-477-346-66	Sequence 66, Appl
3	1058.5	41.1	514	4 US-08-477-089-66	Sequence 66, Appl
4	1058.5	41.1	514	4 US-08-487-072A-66	Sequence 66, Appl
5	383.5	14.9	251	4 US-09-291-170A-13	Sequence 13, Appl
6	383.5	14.9	251	4 US-09-724-884-13	Sequence 13, Appl
7	354	13.8	409	2 US-08-283-917-3	Sequence 3, Appl
8	354	13.8	409	2 US-08-961-716-3	Sequence 3, Appl
9	354	13.8	410	2 US-08-283-917-9	Sequence 9, Appl
10	354	13.8	410	2 US-08-961-716-9	Sequence 9, Appl
11	340.5	13.2	409	1 US-08-190-802A-51	Sequence 51, Appl
12	340.5	13.2	409	3 US-08-477-346-51	Sequence 51, Appl
13	340.5	13.2	409	4 US-08-473-089-51	Sequence 51, Appl
14	340.5	13.2	409	4 US-08-487-072A-51	Sequence 51, Appl
15	323.5	12.6	343	3 US-09-063-743-5	Sequence 5, Appl
16	323.5	12.6	343	4 US-09-590-540-5	Sequence 5, Appl
17	321	12.5	423	4 US-09-013-118-1	Sequence 1, Appl
18	312	12.1	1194	3 US-09-092-508-2	Sequence 2, Appl
19	312	12.1	1194	4 US-09-435-115-2	Sequence 2, Appl
20	312	12.1	1194	4 US-09-069-023-26	Sequence 26, Appl
21	312	12.1	1194	4 US-09-098-310-2	Sequence 2, Appl
22	312	12.1	1205	3 US-09-092-508-16	Sequence 16, Appl
23	312	12.1	1205	4 US-09-435-115-16	Sequence 16, Appl
24	310.5	12.1	640	4 US-09-177-165A-30	Sequence 30, Appl
25	308	12.0	375	3 US-09-063-743-1	Sequence 1, Appl
26	308	12.0	375	4 US-09-590-540-1	Sequence 1, Appl
27	306	11.9	517	1 US-08-190-802A-30	Sequence 30, Appl

28	306	11.9	517	3 US-08-477-346-30	Sequence 30, Appl
29	306	11.9	517	4 US-08-473-089-30	Sequence 30, Appl
30	306	11.9	517	4 US-08-487-072A-30	Sequence 30, Appl
31	305	11.8	514	4 US-09-108-857-2	Sequence 2, Appl
32	302.5	11.8	2627	2 US-08-751-189-3	Sequence 3, Appl
33	302.5	11.8	2627	2 US-09-060-836-3	Sequence 3, Appl
34	302.5	11.8	2627	3 US-09-184-445-3	Sequence 3, Appl
35	297	11.5	587	3 US-08-899-578-2	Sequence 2, Appl
36	294.5	11.4	713	1 US-08-190-802A-63	Sequence 63, Appl
37	294.5	11.4	713	3 US-08-477-346-63	Sequence 63, Appl
38	294.5	11.4	713	4 US-08-473-089-63	Sequence 63, Appl
39	294.5	11.4	713	4 US-08-487-072A-63	Sequence 63, Appl
40	294.5	11.4	713	4 US-09-108-857-3	Sequence 3, Appl
41	294	11.4	253	4 US-09-291-170A-10	Sequence 10, Appl
42	294	11.4	253	4 US-09-724-884-10	Sequence 10, Appl
43	294	11.4	690	4 US-09-291-170A-2	Sequence 2, Appl
44	294	11.4	690	4 US-09-724-884-2	Sequence 2, Appl
45	292.5	11.4	404	4 US-09-184-001-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-190-802A-66  
; Sequence 66, Application US/08190802A  
; Patent No. 5519003  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: P.O. Box 60850  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-0850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,802A  
; FILING DATE: 01-FEB-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 8600-0139  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49  
US-08-190-802A-66

Query Match 41.1%; Score 1058.5; DB 1; Length 514;  
Best Local Similarity 42.5%; Pred. No. 1.1e-97;  
Matches 210; Conservative 97; Mismatches 148; Indels 39; Gaps 12;  
Qy 14 TIQARLVYTGEEAGPPIDLPAGITQQGLICNALL-KNEEATPYLFFV-----GEDEI 66

Db 33 SIKFOALDTGDNVGGALRVPGAISEKQLEELLNQLNGTSDDPVPTFSCTIQGKKASDPV 92  
QY 67 KKSLEDTLDLAS-----VDTENVIDIVYQPAVFKVRPVTRCTSSMPGHAEAVV 115  
Db 93 K-----TIDITDNLSSLIKPGYNSTEDQITLLYTPRAVFKVPVTRSSSAIAGHGSTIL 147  
QY 116 SLNFSP-DGAHLASGSDTTVRLWDLNLTETPHFTCTGKHQWVLCVSWAPDGKRLASGCKA 174  
Db 148 CSAFAPHTSSRMVTGAGDNTARIWDCDTQTPMHTLKGHNWVLCVSWSPDGEVIATGSM 207  
QY 175 GSIIDWPETGQOKGRPLSGHKHINCLAWEPYH--RDPECKRLASASGDGDCRIWDVKL 232  
Db 208 NTIRLWDPKSGQCLGDALRGHSHKWIITSLWEPHILVKGSKPRLASSSKDGTIKIWDTVS 267  
QY 233 GQCLMNIAGHTNAVTAVRWGGAGLIYTSSKDRTVKMWRA-ADGILCRTFSGHAHWVNIA 291  
Db 268 RVCQYTMSTGHTNSVSCVKWGGQLLYSGSHDRTVRVWDINSQGRNCINILKSHAHWNHLS 327  
QY 292 LSTDYVLRGTGPF-HPVKDRSKSHLSLSTEELQESALKRYQAVCP---DEVESLVSCSDDN 347  
Db 328 LSTDYALRIGAFDHTGKKPS-----TPEEAQKKALENYEKICKKNGNSEEMMTASDDY 381  
QY 348 TLYLWRN-NQNKCVERTGHQNVNDVKYSPDVKLIASAFDKSVRLWRASDGQYMATFR 406  
Db 382 TMFLWNPLKSTKPIARMTHGQKLVNHVAFSPDGRYIVSASFSDNSIKLWDRDGRDKFISTFR 441  
QY 407 GHVQAVYTVAWSADSRLLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRV 466  
Db 442 GHIASVYQVAWSSDCRLLVSCSKDTTLKVWDVTRKLSVDLPGIKTKLY-VDWSVDGKRV 500  
QY 467 ASGGKDKVIKLWAY 480  
Db 501 CSGGKDKMVLWTH 514

RESULT 2

US-08-477-346-66  
; Sequence 66, Application US/08477346  
; Patent No. 6262023  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,346  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,072  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; INFORMATION FOR SEQ ID NO: 66:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49  
US-08-477-346-66

Query Match 41.1%; Score 1058.5; DB 3; Length 514;  
Best Local Similarity 42.5%; Pred. No. 1.le-97;  
Matches 210; Conservative 97; Mismatches 148; Indels 39; Gaps 12;  
QY 14 TIQARLVYTGEEAGPPIDLPAGITTOQLGLICNALL-KNEEATPYLFFV-----GEDEI 66  
Db 33 SIKFOALDTGDNVGGALRVPGAISEKQLEELLNQLNGTSDDPVPTFSCTIQGKKASDPV 92  
QY 67 KKSLEDTLDLAS-----VDTENVIDIVYQPAVFKVRPVTRCTSSMPGHAEAVV 115  
Db 93 K-----TIDITDNLSSLIKPGYNSTEDQITLLYTPRAVFKVPVTRSSSAIAGHGSTIL 147  
QY 116 SLNFSP-DGAHLASGSDTTVRLWDLNLTETPHFTCTGKHQWVLCVSWAPDGKRLASGCKA 174  
Db 148 CSAFAPHTSSRMVTGAGDNTARIWDCDTQTPMHTLKGHNWVLCVSWSPDGEVIATGSM 207  
QY 175 GSIIDWPETGQOKGRPLSGHKHINCLAWEPYH--RDPECKRLASASGDGDCRIWDVKL 232  
Db 208 NTIRLWDPKSGQCLGDALRGHSHKWIITSLWEPHILVKGSKPRLASSSKDGTIKIWDTVS 267  
QY 233 GQCLMNIAGHTNAVTAVRWGGAGLIYTSSKDRTVKMWRA-ADGILCRTFSGHAHWVNIA 291  
Db 268 RVCQYTMSTGHTNSVSCVKWGGQLLYSGSHDRTVRVWDINSQGRNCINILKSHAHWNHLS 327  
QY 292 LSTDYVLRGTGPF-HPVKDRSKSHLSLSTEELQESALKRYQAVCP---DEVESLVSCSDDN 347  
Db 328 LSTDYALRIGAFDHTGKKPS-----TPEEAQKKALENYEKICKKNGNSEEMMTASDDY 381  
QY 348 TLYLWRN-NQNKCVERTGHQNVNDVKYSPDVKLIASAFDKSVRLWRASDGQYMATFR 406  
Db 382 TMFLWNPLKSTKPIARMTHGQKLVNHVAFSPDGRYIVSASFSDNSIKLWDRDGRDKFISTFR 441  
QY 407 GHVQAVYTVAWSADSRLLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRV 466  
Db 442 GHIASVYQVAWSSDCRLLVSCSKDTTLKVWDVTRKLSVDLPGIKTKLY-VDWSVDGKRV 500  
QY 467 ASGGKDKVIKLWAY 480  
Db 501 CSGGKDKMVLWTH 514

RESULT 3

US-08-473-089-66  
; Sequence 66, Application US/08473089  
; Patent No. 6342368  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2003, 21:17:36 ; Search time 396 Seconds  
(without alignments)  
1102.932 Million cell updates/sec

Title: US-09-830-980-1  
Perfect score: 2574  
Sequence: 1 MQETDTEQATPHTIQARLV.....PDGSRVASGGKDKVIKLWAY 480

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	2574	100.0	480 23	US-09-830-980-1 Sequence 1, Appli

2	2566	99.7	480	22	US-09-791-537-110509	Sequence 110509,
3	2566	99.7	480	27	US-10-132-744B-10	Sequence 10, Appl
4	2554.5	99.2	481	20	US-09-614-150-5250	Sequence 5250, Ap
5	2554.5	99.2	481	22	US-09-791-537-109339	Sequence 109339,
6	2554.5	99.2	481	32	US-60-191-637-5268	Sequence 5268, Ap
7	2554.5	99.2	481	32	US-60-191-681-4152	Sequence 4152, Ap
8	2534.5	98.5	490	32	US-60-167-217-5359	Sequence 5359, Ap
9	2534.5	98.5	490	32	US-60-173-464-4328	Sequence 4328, Ap
10	1521	59.1	485	27	US-10-132-744A-6	Sequence 6, Appli
11	1521	59.1	485	27	US-10-132-744B-6	Sequence 6, Appli
12	1520	59.1	485	20	US-09-629-469A-11400	Sequence 11400, A
13	1520	59.1	485	22	US-09-791-537-94362	Sequence 94362, A
14	1520	59.1	485	26	US-10-031-660-16	Sequence 16, Appl
15	1520	59.1	485	29	US-10-380-731-574	Sequence 574, App
16	1515	58.9	488	22	US-09-760-469-1273	Sequence 1273, Ap
17	1515	58.9	488	28	US-10-216-583-1273	Sequence 1273, Ap
18	1513	58.8	476	22	US-09-791-537-9973	Sequence 9973, Ap
19	1513	58.8	476	27	US-10-132-744B-9	Sequence 9, Appli
20	1494	58.0	476	23	US-09-830-980-7	Sequence 7, Appli
21	1452	56.4	484	27	US-10-132-744A-2	Sequence 2, Appli
22	1452	56.4	484	27	US-10-132-744B-2	Sequence 2, Appli
23	1393	54.1	480	30	US-10-424-599-249362	Sequence 249362,
24	1375	53.4	499	22	US-09-760-469-1735	Sequence 1735, Ap
25	1375	53.4	499	28	US-10-216-583-1735	Sequence 1735, Ap
26	1354	52.6	557	30	US-10-425-114-72721	Sequence 72721, A
27	1339	52.0	3302	21	US-09-733-089-21333	Sequence 21333, A
28	1339	52.0	3302	23	US-09-816-660-21333	Sequence 21333, A
29	1335.5	51.9	471	19	US-09-513-996A-26337	Sequence 26337, A
30	1335.5	51.9	471	19	US-09-513-996A-58536	Sequence 58536, A
31	1335.5	51.9	471	24	US-09-935-625-6998	Sequence 6998, Ap
32	1335.5	51.9	471	24	US-09-935-625-15984	Sequence 15984, A
33	1335.5	51.9	473	19	US-09-513-996A-26336	Sequence 26336, A
34	1335.5	51.9	473	19	US-09-513-996A-58535	Sequence 58535, A
35	1335.5	51.9	473	24	US-09-935-625-6997	Sequence 6997, Ap
36	1335.5	51.9	473	24	US-09-935-625-15983	Sequence 15983, A
37	1335.5	51.9	490	19	US-09-513-996A-26335	Sequence 26335, A
38	1335.5	51.9	490	19	US-09-513-996A-58534	Sequence 58534, A
39	1335.5	51.9	490	24	US-09-935-625-6996	Sequence 6996, Ap
40	1335.5	51.9	490	24	US-09-935-625-15982	Sequence 15982, A
41	1331	51.7	447	30	US-10-437-963-178074	Sequence 178074,
42	1326.5	51.5	471	20	US-09-620-394B-8342	Sequence 8342, Ap
43	1326.5	51.5	473	20	US-09-620-394B-8341	Sequence 8341, Ap
44	1326.5	51.5	490	20	US-09-620-394B-8340	Sequence 8340, Ap
45	1318.5	51.2	473	22	US-09-791-537-14356	Sequence 14356, A

ALIGNMENTS

RESULT 1  
US-09-830-980-1  
; Sequence 1, Application US/09830980  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, STEPHEN  
; APPLICANT: BOUWMEESTER, ANTONIUS  
; APPLICANT: ROYET, JULIEN  
; TITLE OF INVENTION: REGULATOR OF NOTCH SIGNALING ACTIVITY  
; FILE REFERENCE: 55880(71745)  
; CURRENT APPLICATION NUMBER: US/09/830,980  
; CURRENT FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: PCT/IB99/01891  
; PRIOR FILING DATE: 1999-11-03  
; PRIOR APPLICATION NUMBER: GB 9824045.0  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Drosophila sp.  
US-09-830-980-1

Query Match 100.0%; Score 2574; DB 23; Length 480;



Db 361 ERMTHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGOYMATFRGHVQAVYTVAWSAD 420  
QY 421 SRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 480  
Db 421 SRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 480

RESULT 4  
US-09-614-150-5250  
; Sequence 5250, Application US/09614150  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/09/614,150  
; CURRENT FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5250  
; LENGTH: 481  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-09-614-150-5250

Query Match 99.2%; Score 2554.5; DB 20; Length 481;  
Best Local Similarity 99.6%; Pred. No. 2.8e-240;  
Matches 479; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 MQETDTEQATPHTIQARLVY-TGEEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLF 59  
Db 1 MQETDTEQATPHTIQARLVSDTGEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLF 60  
QY 60 FVGDEIKKSLEDTLDLASVDTENVNIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNF 119  
Db 61 FVGDEIKKSLEDTLDLASVDTENVNIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNF 120  
QY 120 SPDGAHLASGSGDTTVRLWDLNTETPHFTCTGKQWVLCVSWAPDGKRLASGCKAGSIII 179  
Db 121 SPDGAHLASGSGDTTVRLWDLNTETPHFTCTGKQWVLCVSWAPDGKRLASGCKAGSIII 180  
QY 180 WDPETGOQKGRPLSGHKKHINCLAWEPYHRDPYTRCTSSMPGHAEAVVSLNF 239  
Db 181 WDPETGOQKGRPLSGHKKHINCLAWEPYHRDPYTRCTSSMPGHAEAVVSLNF 240  
QY 240 AGHTNAVAVRWGGAGLIYTSSKDRTVKMWRAADGILCRTFSGHAHWVNNIALSTDYVLR 299  
Db 241 AGHTNAVAVRWGGAGLIYTSSKDRTVKMWRAADGILCRTFSGHAHWVNNIALSTDYVLR 300  
QY 300 TGPFFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDDNTLYLWRNNQKNC 359  
Db 301 TGPFFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDDNTLYLWRNNQKNC 360  
QY 360 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGOYMATFRGHVQAVYTVAWSA 419  
Db 361 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGOYMATFRGHVQAVYTVAWSA 420  
QY 420 DSRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 479  
Db 421 DSRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 480  
QY 480 Y 480  
Db 481 Y 481

Db 361 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGOYMATFRGHVQAVYTVAWSA 420  
QY 420 DSRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 479  
Db 421 DSRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 480  
QY 480 Y 480  
Db 481 Y 481  
RESULT 5  
US-09-791-537-109339  
; Sequence 109339, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 109339  
; LENGTH: 481  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-791-537-109339

Query Match 99.2%; Score 2554.5; DB 22; Length 481;  
Best Local Similarity 99.6%; Pred. No. 2.8e-240;  
Matches 479; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 MQETDTEQATPHTIQARLVY-TGEEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLF 59  
Db 1 MQETDTEQATPHTIQARLVSDTGEAGPPIDLPAGITTOOLGLICNALLKNEEATPYLF 60  
QY 60 FVGDEIKKSLEDTLDLASVDTENVNIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNF 119  
Db 61 FVGDEIKKSLEDTLDLASVDTENVNIDIVYQOAVFKVRPVTRCTSSMPGHAEAVVSLNF 120  
QY 120 SPDGAHLASGSGDTTVRLWDLNTETPHFTCTGKQWVLCVSWAPDGKRLASGCKAGSIII 179  
Db 121 SPDGAHLASGSGDTTVRLWDLNTETPHFTCTGKQWVLCVSWAPDGKRLASGCKAGSIII 180  
QY 180 WDPETGOQKGRPLSGHKKHINCLAWEPYHRDPYTRCTSSMPGHAEAVVSLNF 239  
Db 181 WDPETGOQKGRPLSGHKKHINCLAWEPYHRDPYTRCTSSMPGHAEAVVSLNF 240  
QY 240 AGHTNAVAVRWGGAGLIYTSSKDRTVKMWRAADGILCRTFSGHAHWVNNIALSTDYVLR 299  
Db 241 AGHTNAVAVRWGGAGLIYTSSKDRTVKMWRAADGILCRTFSGHAHWVNNIALSTDYVLR 300  
QY 300 TGPFFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDDNTLYLWRNNQKNC 359  
Db 301 TGPFFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDDNTLYLWRNNQKNC 360  
QY 360 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGOYMATFRGHVQAVYTVAWSA 419  
Db 361 VERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDGOYMATFRGHVQAVYTVAWSA 420  
QY 420 DSRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 479  
Db 421 DSRLIVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLWAY 480  
QY 480 Y 480  
Db 481 Y 481







Qy 60 FVGEDEIKKSLEDTLDLASVDTENVIDIVYQPOAVFKVRPVRTCTSSMPGHAEAVVSLNF 119  
Db 61 FVGEDEIKKSLEDTLDLASVDTENVIDIVYQPOAVFKVRPVRTCTSSMPGHAEAVVSLNF 120  
Qy 120 SPDGAHLASGSGDTTVRLWDLNLTETPHFTCTGHKQWVLCVSWAPDGRKRLASGCKAGSI 179  
Db 121 SPDGAHLASGSGDTTVRLWDLNLTETPHFTCTGHKQWVLCVSWAPDGRKRLASGCKAGSI 180  
Qy 180 WDPETGOQKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 239  
Db 181 WDPETGOQKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 240  
Qy 240 AGHTNAVTAVRWGGAGLIYTSSKDRVTVMWRAADGILCRTPSGHAHWVNNIALSTDYVLR 299  
Db 241 AGHTNAVTAVRWGGAGLIYTSSKDRVTVMWRAADGILCRTPSGHAHWVNNIALSTDYVLR 300  
Qy 300 TGPFPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQNK 359  
Db 301 TGPFPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQNK 360  
Qy 360 VERMTGHQNVNDVKYSPDVKLIASAFDKSVRLWRASDGOVMATFRGHVQAVYTVAWSA 419  
Db 361 VERMTGHQNVNDVKYSPDVKLIASAFDKSVRLWRASDGOVMATFRGHVQAVYTVAWSA 420  
Qy 420 DSRLIVSGSKDSTLKWMSVQTKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLM 478  
Db 421 DSRLIVSGSKDSTLKWMSVQTKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLY 479

RESULT 9

US-60-173-464-4328  
; Sequence 4328, Application US/60173464  
; GENERAL INFORMATION:  
; APPLICANT: Li, Peter W.D.  
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000173  
; CURRENT FILING DATE: 1999-12-29  
; NUMBER OF SEQ ID NOS: 30269  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4328  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Drosophila  
US-60-173-464-4328

Query Match 98.5%; Score 2534.5; DB 32; Length 490;  
Best Local Similarity 99.4%; Pred. No. 2.6e-238;  
Matches 476; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 MQETDTEQEAETHPTIARLVY-TGEEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLF 59  
Db 1 MQETDTEQEAETHPTIARLVSDTGEEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLF 60  
Qy 60 FVGEDEIKKSLEDTLDLASVDTENVIDIVYQPOAVFKVRPVRTCTSSMPGHAEAVVSLNF 119  
Db 61 FVGEDEIKKSLEDTLDLASVDTENVIDIVYQPOAVFKVRPVRTCTSSMPGHAEAVVSLNF 120  
Qy 120 SPDGAHLASGSGDTTVRLWDLNLTETPHFTCTGHKQWVLCVSWAPDGRKRLASGCKAGSI 179  
Db 121 SPDGAHLASGSGDTTVRLWDLNLTETPHFTCTGHKQWVLCVSWAPDGRKRLASGCKAGSI 180  
Qy 180 WDPETGOQKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 239  
Db 181 WDPETGOQKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNI 240  
Qy 240 AGHTNAVTAVRWGGAGLIYTSSKDRVTVMWRAADGILCRTPSGHAHWVNNIALSTDYVLR 299  
Db 241 AGHTNAVTAVRWGGAGLIYTSSKDRVTVMWRAADGILCRTPSGHAHWVNNIALSTDYVLR 300

Qy 300 TGPFPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQNK 359  
Db 301 TGPFPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQNK 360  
Qy 360 VERMTGHQNVNDVKYSPDVKLIASAFDKSVRLWRASDGOVMATFRGHVQAVYTVAWSA 419  
Db 361 VERMTGHQNVNDVKYSPDVKLIASAFDKSVRLWRASDGOVMATFRGHVQAVYTVAWSA 420  
Qy 420 DSRLIVSGSKDSTLKWMSVQTKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLM 478  
Db 421 DSRLIVSGSKDSTLKWMSVQTKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKLY 479

RESULT 10  
US-10-132-744A-6  
; Sequence 6, Application US/10132744A  
; GENERAL INFORMATION:  
; APPLICANT: Utku, Nalan  
; TITLE OF INVENTION: Novel genes Tzap7/A, Tzap7/B and Tzap7 involved in T cell  
; TITLE OF INVENTION: activation and uses thereof  
; FILE REFERENCE: Utku-4 CON  
; CURRENT APPLICATION NUMBER: US/10/132,744A  
; PRIOR FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: PCT/EP00/10670  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: 60/185,016  
; PRIOR FILING DATE: 2000-02-25  
; PRIOR APPLICATION NUMBER: 60/162,675  
; PRIOR FILING DATE: 1999-11-01  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: miscellaneous feature  
; LOCATION: 379  
; OTHER INFORMATION: variable amino acid  
US-10-132-744A-6

Query Match 59.1%; Score 1521; DB 27; Length 485;  
Best Local Similarity 58.6%; Pred. No. 4.8e-139;  
Matches 279; Conservative 73; Mismatches 118; Indels 6; Gaps 3;

Qy 9 EATPHTIQARLVYTGEAA---GPPIDLPAGITTOQLGLICNALLKNEEATPYLFFVGED 64  
Db 8 EAVARDVQRLLVQFQDEGGQLGSPFDPVDITPDRLLQVLCNALLAQEDPLPLAFFVHDA 67  
Qy 65 EIKKSLEDTLDLASVDTENVIDIVYQPOAVFKVRPVRTCTSSMPGHAEAVVSLNFPDGA 124  
Db 68 EIVSSLGKTLESQAVETEKVLDIIYQQAIFRVRVAVTRCTSSLEGSEAVISVAFSPTGK 127  
Qy 125 HLASGSGDTTVRLWDLNLTETPHFTCTGHKQWVLCVSWAPDGRKRLASGCKAGSI 184  
Db 128 YLASGSGDTTVRFWDLSTETPHFTCKGHRHWVLSISWSPDGKKLASGCKNGQILLWDPST 187  
Qy 185 GQOKGRPLSGHKHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGQCLMNIAGHTN 244  
Db 188 GKQVGRTLAGHSKWITGLSWEPLHANPECRYVASSSKDGSVRWDITTAGRCERILTGHTQ 247  
Qy 245 AVTAVRWGGAGLIYTSSKDRVTVMWRAADGILCRTPSGHAHWVNNIALSTDYVLRTPFH 304  
Db 248 SVTCLRWGGDGLLYSASQDRTIKVMRAHDGVLCTLQGHGHWVNTMALSTDYALRTGAPE 307  
Qy 305 PVK-DRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQNK-KCVER 362  
Db 308 PAASVNPQDLQSGSLQELKERALSRYNLVRGQGPRLVSGSDDFLFLWSPAEDKKPLTR 367  
Qy 363 MTGHQNVNDVKYSPDVKLIASAFDKSVRLWRASDGOVMATFRGHVQAVYTVAWSADSR 422  
Db 368 MTGHQALINQVXFSPDSRIVASAFDKSIKLDWGRGTGKYLASLRGHVAAVYQIAWSADSR 427

[illegible]

```

; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11400
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-629-469A-11400

Query Match      59.1%; Score 1520; DB 20; Length 485;
Best Local Similarity 58.6%; Pred. No. 6e-139;
Matches 279; Conservative 73; Mismatches 118; Indels 6; Gaps 3

QY      9  EATPHTIARLVYTGEEA----GPPIDLPAGITTCQLGLICNALLKNEEATPYLPFVGED 64
      ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      8  EAVARDVQRLLVQFQDEGGQLLGSFPDVPVDITPDRLLQVLCNALLAQEDPLPLAFVHDA 67

QY      65  EIKKSLDITDLASVDTENVIDIVYQPAVFKVRPVTCTSSMPGHAEAVVSLNFSPDGA 124
      ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      68  EIVSSILGKTLESQAVETEKVLDIYQPAIFRVRAVTRCTSSLEGHSEAVISVAFSPTGK 127

QY      125  HLASGSDTTVRLWDLNLTETPHFTCTGHKQWVLCVSWAPDGKRLASGCKAGSIIDWPET 184
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      128  YLASGSDTTVRFWDLSTETPHFTCKGRRHWVLSISWSPDGKKLASGCKNGQILLWDPST 187

QY      185  GQKGRPLSGHKXGHINCLAWEPYHRDPECRKLASASGDCDCRIWDVKLGQCLMNIAGHTN 244
      |:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      188  GKQVGRTLAGHSKWI TGLSWEPLHANPECRYVASSSKDGSVRIWDTTAGRCERILTGTQ 247

QY      245  AVTAVRWCGAGLIYTSKDRITVQWRAADGILCRFESGHAHWNNIALSTDYVLRGTGPFH 304
      : || :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      248  SVTCLRWGGDGLLYSASQDRTIKVWRAHDGVLCTLQGHGHVWNTMALSTDYALRTGAF 307

QY      305  PVK-DRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDDNLTLYLRNNQN-KCVER 362
      |:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      308  PAEASVNPQDLQGSLOELKELRALSRYNLVRGQGPRLVSGSDDFTLFLWSPAEDKKPLTR 367

QY      363  MTGHQNVVNDVKYSPDVKLIASASFDKSVRLWRASDQYMATFRGHVQAVYTVAWSADSR 422
      ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      368  MTGHQALINQVLFSPDSRIVASASFDKSIKLWDGRTGKYLASLRGHVAAVYQIAWSADSR 427

QY      423  LIVSGSKDSTLLKWSVQTKKLAQELPGHADEVFGVDWADPGSRVASGGKQKVIKLW 478
      |:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      428  LLVSGSDSTLLKVDWKAQKLANDLPGHADEVYAVDWSPDGQRVASGGKQKCLRIV 483

```

```

; APPLICANT: LU, Dyung Aina M.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: GTP-BINDING ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0714 PCT
; CURRENT APPLICATION NUMBER: US/10/031,660
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/144,595; 60/150,460; 60/159,849
; PRIOR FILING DATE: 1999-07-19; 1999-08-23; 1999-10-15
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2216640CD1
US-10-031-660-16

Query Match      59.1%; Score 1520; DB 26; Length 485;
Best Local Similarity 58.6%; Pred. No. 6e-139;
Matches 279; Conservative 73; Mismatches 118; Indels 6; Gaps 3;

QY 9 EATPHTIQARLVYTGEEA- - - GPIDLPAGITTOQLGLICNALLKNEEATPYLFFVGED 64
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 8 EAVARDVQRLVQFQDEGGQLGSPFDVPVDITPDRLLQVNCNALLAQEDPLPLAFFVHDA 67
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 65 EIKKSLSDTLDLASVDTENVIDIVYQPAVFKVRPVTRCTSSMPGHAEVVSLSNFSPDGA 124
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 68 EIVSSLGKTLESQAVETEKVLDIIYQPAIFRVRAVTRCTSSLEGHSEAVISVAFSPTGK 127
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 125 HLASGSGDTTVRLWDLNTETPHFTCTGHKQWVLCVSWAPDGKRLASGCKAGSI IWDPET 184
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 128 YLASGSGDTTVRFWDLSTETPHFTCKGHRHWVLSISWSPDGKKLASGCKNQGILLWDPST 187
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 185 GQKGRPLSGHKGHINCLAWEPYHRDPECRKLASASGDGDCRIWDVKLGOC LMNIAGHTN 244
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 188 GKQVGRTLAGHSKWI TGLSWEPLHANPECRYVASSXKOGSVRIWDTTAGRCERILTGHTQ 247
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 245 AVTAVRWGGAGLIYTSSKDRTVKMWRAADGILCRTFSGHAHWVNNTALSTDYVLRGTGPFH 304
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 248 SVTCLRWGGDGLLYSASQDRTIKVRRAHDGVL CRTLOQGHGWVNTMALSTDYALRTGAFE 307
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 305 PVK-DRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDDNTLYLWRNNQN-KCVER 362
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 308 PAEASVNPQDLQGLQELKERALSRYNLVRGGPERLVSGSDDFTFLWSPAEDEKKPLTR 367
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 363 MTGHQNVVNDVKYSPDVKLIASAFDKSVRLWRASDGQYMATFRGHVQAVYTVAWSADSR 422
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 368 MTGHQALINQVLFSPDSRIVASASFDKSIKLDWGR TGKYLASLRGHVAAVYQIAWSADSR 427
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

QY 423 LTVSGSKDSTLKVWSVQTKKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKWLW 478
   | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 428 LLVSGSSDSTLKVWDVKAQKLAMDLPGHADDEVYAVDWSPDGQORVASGGKDKKCLRIMW 483
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 15
US-10-380-731-574
; Sequence 574, Application US/10380731
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-114
; CURRENT APPLICATION NUMBER: US/10/380,731
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/659,671
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 888
; SOFTWARE: Custom
; SEQ ID NO 574
; LENGTH: 485
; TYPE: PRT

```

; ORGANISM: Homo sapiens  
US-10-380-731-574

Query Match		59.1%	Score 1520;	DB 29;	Length 485;
Best Local Similarity		58.6%	Pred. No. 6e-139;		
Matches	279;	Conservative	73;	Mismatches	118;
				Indels	6;
				Gaps	3;
QY	9	EATPHITQARLVYTGEEA----	GPPIDLPAGITTTQQLGLICNALLKNEEATPYLFFVGED	64	
Db	8	EAVARDVQRLLVQFQDEGGQLLGSPPFDPVDITPDRQLQVNCNALLAQEDPLPLAFFVHDA	67		
QY	65	EIKKSLEDTLDLASVDTENVIDIYQQAQVFKVRPVTTRCTSSMPGHAEAVVSLNFSPDGA	124		
Db	68	EIVSSLGKTLESQAVETEKVLDIYQQAIFRVRVAVTRCTSSLEGHSEAVISVAFSPTGK	127		
QY	125	HLASGSGDTTVRLWDLNTETPHFTCTGHKQWVLCVSWAPDGGKRLASGCKAGSIILWDPET	184		
Db	128	YLASGSGDTTVRFWDLSTETPHFTCKGHRHWVLSISWSPDGKKLASGCKNGQILLWDPST	187		
QY	185	GQKGRPLSGHKHINCLAWEPYHRDPECRKCLASASGDDCRIWDVKLGQCLMNIAGHTN	244		
Db	188	GKQVGRITLAGHSHKMITGLSWEPLHANPECRYVASSSKDGSVRIMDTTAGRCERILTGTQ	247		
QY	245	AVTAVRWGGAGLIYTSKDRTVKMWRAADGILCRTFSGHAHWVNNIALSTDYVLRITGPFH	304		
Db	248	SVTCLRWGGDGLLYSASQDRITIKVWRAHDGVLCRTLQGHGHVNTMALSTDYALRTGAFE	307		
QY	305	PVK-DRSKSHLSLSTEELQESALKRYQAQVCPDEVESLVSCSDDNTLYLWRNNON-KCVER	362		
Db	308	PAEASVNPQDLQGSLOELKERALSRYNLVRGQGPRLVSGSDDFTLFLWSPAEDKKPLTR	367		
QY	363	MTGHQNVVNDVKYSPDVKLIASAFDKSVRLWRASDGQYNATFRGHVQAVYTVAWSADSR	422		
Db	368	MTGHQALINQVLFSPDSRIVASAFDKSIKLDWGRITGKYLASLRGHVAAVYQIAWSADSR	427		
QY	423	LIVSGSKDSTLKVMSVQTKLAQELPGHADEVFGVDWAPDGSRVASGGKDKVIKILW	478		
Db	428	LLVSGSSDSTLKVWDVKAQKLANDLPGHADDEVYAVDWSPDQQRVASGGKDKCLRIW	483		

Search completed: October 10, 2003, 21:27:56  
Job time : 398 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2003, 21:06:46 ; Search time 83 Seconds  
(without alignments)  
917.937 Million cell updates/sec

Title: US-09-830-980-1  
Perfect score: 2574  
Sequence: 1 MQETDTEQATPHTIQARLV.....PDGSRVASGGKDKVIKLWAY 480

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03: \*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: \*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: \*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: \*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: \*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: \*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: \*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: \*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: \*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: \*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: \*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: \*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: \*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: \*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: \*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: \*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: \*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: \*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: \*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: \*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: \*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: \*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: \*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: \*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	2558.5	99.4	479	21 AAY79678	Drosophila Notchle
2	2554.5	99.2	481	22 ABB59486	Drosophila melanog
3	1521	59.1	485	22 AAB68284	Amino acid sequenc
4	1520	59.1	485	22 AAB92844	Human protein sequ
5	1520	59.1	485	22 AAB68516	Human GTP-binding
6	1520	59.1	485	23 ABB97306	Novel human protei
7	1452	56.4	484	22 AAB68282	Amino acid sequenc
8	1335.5	51.9	471	21 AAG23141	Arabidopsis thalia
9	1335.5	51.9	471	21 AAG46521	Arabidopsis thalia

10	1335.5	51.9	473	21 AAG23140	Arabidopsis thalia
11	1335.5	51.9	473	21 AAG46520	Arabidopsis thalia
12	1335.5	51.9	490	21 AAG23139	Arabidopsis thalia
13	1335.5	51.9	490	21 AAG46519	Arabidopsis thalia
14	1078.5	41.9	515	24 ABJ26155	Aspergillus fumiga
15	1058.5	41.1	514	16 AAR85881	WD-40 domain-contg
16	1001	38.9	520	23 ABP73383	Candida albicans e
17	819	31.8	261	22 AAB68283	Amino acid sequenc
18	783.5	30.4	435	24 ABJ25555	Aspergillus fumiga
19	419	16.3	317	21 AAB63186	Human secreted pro
20	419	16.3	334	22 AAB68529	Human GTP-binding
21	419	16.3	334	23 ABB97345	Novel human protei
22	415	16.1	361	22 ABB68576	Drosophila melanog
23	412.5	16.0	700	22 ABB60376	Drosophila melanog
24	399	15.5	330	22 AAB93659	Human protein sequ
25	398.5	15.5	341	21 AAG38744	Arabidopsis thalia
26	389	15.1	411	22 ABB62260	Drosophila melanog
27	372.5	14.5	514	22 AAB95225	Human protein sequ
28	372	14.5	323	21 AAB63185	Gene 3 human secre
29	371	14.4	542	23 ABP41760	Human ovarian anti
30	362	14.1	696	24 ABJ25848	Aspergillus fumiga
31	362	14.1	696	24 ABJ26448	Aspergillus fumiga
32	355	13.8	333	21 AAG05554	Arabidopsis thalia
33	354	13.8	409	16 AAR70002	OPD8 45 kDa subuni
34	354	13.8	410	16 AAR70005	OPDE 45 kDa subuni
35	348	13.5	450	24 ABB99407	Amino acid sequenc
36	348	13.5	478	22 AAM93784	Human polypeptide,
37	348	13.5	478	24 ABB99402	Amino acid sequenc
38	348	13.5	521	22 ABB10141	Human cDNA SEQ ID
39	348	13.5	521	23 ABP66728	Human polypeptide
40	344	13.4	159	20 AAY10919	Amino acid sequenc
41	343	13.3	407	22 AAM93675	Human polypeptide,
42	340.5	13.2	409	16 AAR85868	WD-40 domain-contg
43	335	13.0	358	22 ABB65223	Drosophila melanog
44	331	12.9	584	22 ABB21351	Novel human diagno
45	330	12.8	423	22 AAB94261	Human protein sequ

ALIGNMENTS

RESULT 1						
AAAY79678						
ID	AAAY79678	standard; Protein; 479 AA.				
XX						
AC	AAAY79678;					
XX						
DT	29-AUG-2000	(first entry)				
XX						
DE	Drosophila Notchless protein.					
XX						
KW	Notchless; Nle gene; Notch; signalling; neurodegenerative disease;					
KW	cancer; diagnosis; cytostatic; neuroprotective; therapy.					
XX						
OS	Drosophila melanogaster.					
XX						
FH	Key	Location/Qualifiers				
FT	Domain	27..105				
FT			/note= "Nle domain"			
FT	Misc-difference 67					
FT			/note= "sequence deduced from nucleotide sequence			
FT			has an additional Lys residue between			
FT			Lys-67 and Ser-68"			
FT	Misc-difference 282					
FT			/note= "encoded by CAA"			
XX						
PN	WO200026364-A1.					
XX						
PD	11-MAY-2000.					
XX						
PF	03-NOV-1999;		99WO-IB01891.			
XX						
PR	03-NOV-1998;		98GB-0024045.			

Not a protein

XX (EUMO-) EURO MOLECULAR BIOLOGY LAB.  
PA Cohen S, Bouwmeester A, Royet J;  
XX WPI; 2000-365613/31.  
DR N-PSDB; AAA27739.  
XX Novel Notchless protein and nucleic acids encoding them useful for  
PT treating and preventing cancer and neurodegenerative diseases -  
PT  
XX Claim 1; Page 44; 52pp; English.  
XX The present sequence is that of Notchless, a novel protein of  
CC Drosophila. Notchless was identified in a screen for dominant  
CC modifiers of a Notch mutant phenotype in the Drosophila wing. The  
CC mutant dominantly suppressed the wing notching phenotype of  
CC notchoid mutations, and the Notchless protein was shown to bind to  
CC the cytoplasmic domain of Notch. Notchless modified Notch  
CC signalling activity in a variety of Notch-dependent signalling  
CC process in both Drosophila and Xenopus embryos. The Notchless  
CC protein has a novel highly conserved N-terminal domain followed by  
CC 9 WD40 repeats. Notchless, and nucleic acids encoding it, can  
CC be used in methods for the diagnosis and therapy of certain diseases,  
CC particularly cancer and neurodegenerative diseases (claimed). A  
CC Notchless mutant in a sensitised Notch genetic background is used  
CC in a claimed method for identifying compounds capable of modifying  
CC the levels of expression or activity of a Notch protein.  
XX  
SQ Sequence 479 AA;  
Query Match 99.4%; Score 2558.5; DB 21; Length 479;  
Best Local Similarity 99.8%; Pred. No. 7e-233;  
Matches 479; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 MQETDTEQATPHTIQARLVYTGEEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLFF 60  
DB 1 MQETDTEQATPHTIQARLVYTGEEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLFF 60  
QY 61 VGEDEIKKSLEDTLDLASVDTEENVIDIVYQPAVFKVRPVTTRCTSSMPGHAEAVVSLNFS 120  
DB 61 VGEDEIKKSLEDTLDLASVDTEENVIDIVYQPAVFKVRPVTTRCTSSMPGHAEAVVSLNFS 119  
QY 121 PDGAHLASGSGDITVRLWDLNLTETPHFTCTGKQWVLCVSWAPDGRKRLASGCKAGSI 180  
DB 120 PDGAHLASGSGDITVRLWDLNLTETPHFTCTGKQWVLCVSWAPDGRKRLASGCKAGSI 179  
QY 181 DPETGQKGRPLSGHKHINCLAWEPYHRDPECKRLASASGDGDCRIWDVKGQCLMNI 240  
DB 180 DPETGQKGRPLSGHKHINCLAWEPYHRDPECKRLASASGDGDCRIWDVKGQCLMNI 239  
QY 241 GHTNAVTAVRWGGAGLIYTSKDRVTVMWRAADGILCRFTSGHAHWVNNIALSTDYVLR 300  
DB 240 GHTNAVTAVRWGGAGLIYTSKDRVTVMWRAADGILCRFTSGHAHWVNNIALSTDYVLR 299  
QY 301 GPFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQKCV 360  
DB 300 GPFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQKCV 359  
QY 361 ERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDQGYMATFRGHVQAVYTVAWSAD 420  
DB 360 ERMTGHQNVNDVKYSPDVKLIASASFDKSVRLWRASDQGYMATFRGHVQAVYTVAWSAD 419  
QY 421 SRLIVSGSKDSTLKWSVQTKLAQELPGHADEVFGVDWAPDGRSVASGDKVIKLWAY 480  
DB 420 SRLIVSGSKDSTLKWSVQTKLAQELPGHADEVFGVDWAPDGRSVASGDKVIKLWAY 479

XX 26-MAR-2002 (first entry)  
XX Drosophila melanogaster polypeptide SEQ ID NO 5250.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.  
XX Drosophila melanogaster.  
XX WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US09231.  
XX 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX N-PSDB; ABL03589.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure; SEQ ID NO 5250; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
XX The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 481 AA;  
Query Match 99.2%; Score 2554.5; DB 22; Length 481;  
Best Local Similarity 99.6%; Pred. No. 1.7e-232;  
Matches 479; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 MQETDTEQATPHTIQARLVYTGEEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLFF 59  
DB 1 MQETDTEQATPHTIQARLVYTGEEAGPPIDLPAGITTOQLGLICNALLKNEEATPYLFF 60  
QY 60 FVGEDEIKKSLEDTLDLASVDTEENVIDIVYQPAVFKVRPVTTRCTSSMPGHAEAVVSLNFS 119  
DB 61 FVGEDEIKKSLEDTLDLASVDTEENVIDIVYQPAVFKVRPVTTRCTSSMPGHAEAVVSLNFS 120  
QY 120 SPDGAHLASGSGDITVRLWDLNLTETPHFTCTGKQWVLCVSWAPDGRKRLASGCKAGSI 179  
DB 121 SPDGAHLASGSGDITVRLWDLNLTETPHFTCTGKQWVLCVSWAPDGRKRLASGCKAGSI 180  
QY 180 WDPETGQKGRPLSGHKHINCLAWEPYHRDPECKRLASASGDGDCRIWDVKGQCLMNI 239  
DB 181 WDPETGQKGRPLSGHKHINCLAWEPYHRDPECKRLASASGDGDCRIWDVKGQCLMNI 240  
QY 240 AGHTNAVTAVRWGGAGLIYTSKDRVTVMWRAADGILCRFTSGHAHWVNNIALSTDYVLR 299  
DB 241 AGHTNAVTAVRWGGAGLIYTSKDRVTVMWRAADGILCRFTSGHAHWVNNIALSTDYVLR 300  
QY 300 TGFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQKNC 359  
DB 301 TGFHPVKDRSKSHLSLSTEELQESALKRYQAVCPDEVESLVSCSDNTLYLWRNNQKNC 360